The "unreasonable effectiveness of mathematics" in biology and the fallacy from complexity

Gregor P. Greslehner

Author's accepted manuscript, published in: Gabriele M. Mras, Paul Weingartner & Bernhard Ritter (eds.) *Contributions of the Austrian Ludwig Wittgenstein Society Volume XXVI. Philosophy of Logic and Mathematics. Contributions of the 41st International Wittgenstein Symposium*, Austrian Ludwig Wittgenstein Society: 70–72, 2018.

University of Salzburg, Austria
gregor.greslehner@gmail.com

Abstract

A frequently put forward argument claims that biological systems are too complex for mathematical methods to be fruitfully applied. I argue that this argument from complexity is a fallacy. To the contrary, it is exactly the complexity of biological systems which calls for the use of mathematical methods. While some research strategies in molecular biology used to be less accessible for mathematical analysis, the emergence of systems biology as a scientific discipline is the most recent example of successful and effective applications of mathematical methods in biology. In today's scientific practice, there is evidence for the "unreasonable effectiveness of mathematics" in biology, differing from traditional mathematical biology, thus giving new support to a notion that still cannot be taken for granted.

1 Introduction

Contrary to the "unreasonable effectiveness of mathematics in the natural sciences" (Wigner, 1960), the usefulness of mathematical methods in areas other than physics – biology, in particular – have often been viewed with skepticism. One of the main reasons for this misconception, I claim, is a fallacious argument about the complexity of biological systems.

Eugene P. Wigner describes his wonder about the applicability of mathematics in physics as follows:

> "The miracle of the appropriateness of the language of mathematics for the formulation of the laws of physics is a wonderful gift which we neither understand nor deserve." (Wigner, 1960, 14)

Despite considering its applicability to be a miracle, Wigner is quite optimistic when it comes to physics and the inanimate world, to which he explicitly restricts his discussion (Wigner, 1960, 3, 5, 11), considering the laws, generalizations, and regularities in nature, which can be captured in a mathematical language in physics. However, when it comes to complex phenomena in the realm of biology, he is more cautious:

> "A much more difficult and confusing situation would arise if we could, some day, establish a theory of the phenomena of consciousness, or of biology, which would be as coherent and convincing as our present theories of the inanimate world." (Wigner, 1960, 13)

In a similar vein, Israel M. Gelfand referred to "the unreasonable ineffectiveness of mathematics in biology" (Lesk, 2000, 29). Why should mathematics be less effective in biology? A common conception is that biology's complexity is responsible for this.

2 The fallacy from complexity

While it is rarely defended in print, a common argument against the usefulness of mathematical methods in biology is constructed from the complexity of biological systems. Jacob T. Schwartz, for example, puts it like this:

> "Mathematics is able to deal successfully only with the simplest of situations, more precisely, with a complex situation only to the extent that rare good fortune makes this complex situation hinge upon a few dominant simple factors." (Schwartz, 1992, 21–22)

I call this the *fallacy from complexity*: because biology is complex, mathematics is not helpful in biology. This argument from complexity can be reconstructed as follows:

1. Mathematics can only be useful in simple cases.
2. Biology is not simple.
3. Thus, mathematics is not useful in biology.

Although logically valid, my main objection against this argument is denying the truth of the first premise. Mathematics is not restricted to simple cases. It is capable of dealing with highly complex situations, depending on multiple factors of which the dominant ones cannot be easily identified. Even if it were true that mathematics is only useful in situations depending a few dominant factors, does the scarcity of finding such factors rule out any applicability of mathematics? No, since identifying the relevant variables in a complex system is a mathematical task in the first place. *A fortiori*, if there are no simple dominant factors to be easily identified, mathematical tools for analyzing data, finding patterns, and building models that can then be tested empirically, are required.

The argument from complexity is mistaken, I think, insofar as exactly in those cases where it is difficult to identify the “dominant factors” that mathematics, data science, and computational methods are the most expedient tools one could use. What would be the alternative? Using intuition, adding more facts to the description of the system, hoping that in the future a simple model might emerge? I think that it is in exactly those situations that the mathematical art of model building and data analysis is the route to go. This holds especially for complex biological systems.

> “Our intuitive understanding of phenomena cannot cope with the overwhelming complexity of nonlinear interactions among cellular components. The use of mathematical models is then an indispensable tool to tackle this complexity systematically.” (Boogerd et al., 2007, 15)

We need mathematics in situations that are not simple. If things are simple, one could argue that there would be no need to bring in mathematics. But as soon as rigorous analysis is required to make sense of more complex cases, it is an indispensable tool. Thus, contrary to the first premise of the fallacy from complexity, mathematics is useful foremost in cases that are not simple. In fact, the effectiveness of mathematics becomes apparent particularly in complex situations. Researchers in systems biology are perfectly aware of this when discussing the role of mathematics in their field:

> “A defining feature of systems biology is the role that mathematical modelling plays. [. . .] for mathematical modelling to be employed we require a certain level of complexity that is necessary to convince the nonmathematician of its usefulness.” (Wolkenhauer and Ullah, 2007, 165)

A different objection against the argument from complexity can be made regarding the second premise. One way to read the statement that biology is not simple is that there are no fundamental principles for biologists to discover. This is not the case. However, these principles are different from those in physics: exceptions exist, boundary conditions are more important, and models are less elegant due to the large number of relevant factors. But still, for understanding complex systems, simple models do provide deep insight in biology. Put this way, fundamental principles in biology can be simple.

3 Why strive for simplicity in biology?

Another issue arising from the debate concerning complexity and simplicity in biological systems is the question why simplicity should be aimed for in biology. Looking for simplicity and elegance is a common motive among mathematicians. While usually a helpful guideline in deciding between competing hypotheses, it is not necessarily expedient to expect such features in biological systems that have been shaped by natural selection.

> "The simplicity of mechanisms that serves as Occam's razor in the decision between competing theories in physics is of comparatively lower real value in biology. Functionality and fitness and empirical facts rule over simplicity. The actual mechanisms in systems biology may be more complex than possible because of coselection for other purposes in evolutionary optimization, because evolution may have led to systems that are optimal locally but not globally, and because simplicity in human eyes may be complex in systems biology terms (and vice versa)." (Westerhoff and Kell, 2007, 47, original italics)

The detailed study of biological entities and their activities typically requires years of experimental work in the lab. For the lab researcher, the use and development of sophisticated mathematical methods and models is less attractive. Mathematical methods do play a role in data acquisition and processing, in how to represent and interpret them statistically, e.g., with significance tests. But, as it turns out, a lot of mistakes take place in how these methods are used in biology (May, 2004). The way in which experimental data is represented, stored, processed, and accessed are interesting problems. How to make sense of the data is crucial as the amount of data accumulating in biological databases continues to grow. There is a steadily increasing gap between the data sets that are available and the biological knowledge that can be extracted from them. Merely collecting the data is not enough. In order to interpret and make sense of these data, other mathematical approaches than those for data acquisition are required.

According to a statement attributed to Ernest Rutherford, "all science is either physics or stamp collecting". Compared to physics, biology appears to be more about finding empirical details than working out fundamental laws of nature. If we take Rutherford's quote as a distinction between theory building and fact finding, biology puts more emphasis on the latter and used to be rather descriptive and qualitative. But there is no need to disqualify it from being a science or to contemptuously call it "stamp collecting". Contemporary biology is putting more emphasis on formal representations of complex systems. There even is a scientific discipline devoted explicitly to studying complex biological systems, putting mathematical approaches center stage: systems biology.

4 Systems biology

If mathematics is viewed primarily as a kind of language, then it may seem obvious that in biology, where one needs to describe a plethora of details, the benefit from a "mathematical shorthand" in addition to established nomenclature would not make working in biology easier – at least for descriptive purposes, which have been predominant in biology so far. For big data biology, however, a solely descriptive approach is no longer practicable. In many cases applying mathematical formalism results in novel insights into the structure of biological systems that cannot be obtained simply by the descriptive data collecting approach. There seems to be a growing awareness that although data collecting remains an important task for biology in the future, and fantastic novel techniques broaden the range of biotechnology, formal approaches to make meaningful use of the overwhelming amount of data already collected are needed for a systematic understanding of the working mechanisms of biological systems. There cannot be a biology without overarching models and theories, the appropriate language for which is mathematics.

One of the most important aims is to develop a more formal language to express quantitative relationships in biological models, "[t]he need for a new language in functional biology, more adequate than the diagrams and non-formal language usually found in molecular and cell biology" (Braillard, 2015, 353). When it comes to mapping the structural and sequence data of biological components to their molecular and physiological functions, a regimented vocabulary is a necessary tool. Steps in this direction can be seen in the Gene Ontology project (Ashburner et al., 2000) or the Systems Biology Markup Language (Hucka et al., 2003). This is where mathematical language and formalism have a major impact on biological thinking.

In dealing with complex systems, this is exactly where the art of mathematical modeling is important: to find the right balance between too little and too much details, identifying the relevant factors, and pattern finding. This is what makes drawing the conclusion that mathematics is not useful from the complexity of biological systems a fallacy. To the contrary, *because* biology is complex, it needs mathematics. Faced with complex biological data, mathematics is our best ally.

The traditional skepticism about the effectiveness of mathematics in biology has been rooted to some degree in the scientific situation of biology. The key players (molecules, genes) have mostly been identified in the last decades. What is needed now is a quantitative understanding of how these components interact to bring about biological phenomena on higher levels of organization.

Mathematical tools are fruitfully applied in systems biology, like analysis of networks with graph theory and dynamical systems theory, see for example (Alon, 2007). Many of the tools it would take to advance theory building and modeling in biology do not need to be developed from scratch but could be achieved with more interdisciplinary cooperation between biologists and mathematicians, data scientists, statisticians – and philosophers.

> "A new comprehensive theoretical biology understood as a merger of mathematical biology, bioinformatics and theoretical systems biology, is still far away from the state of perfection in theoretical chemistry but it is making fast progress and together with the spectacular achievements in experimental techniques it sets the stage for a new understanding of biology." (Schuster, 2011, 8)

5 Conclusion

Mathematics strives for simplicity and elegance, whereas the "messy" world of biology abounds with intricate descriptions of molecular details. This is no sign of mathematics being ineffective in biology but does justice to the rich diversity of biological systems. While molecular biology has identified the key components in the past, systems biology is now addressing the interactions and dynamics of these components with mathematical tools. Mathematics is providing the adequate means for dealing with complex biological systems on several levels of organization, thus being "unreasonably effective" in biology – or rather, quite reasonably so.

Acknowledgments

Thanks to Laurenz Hudetz, Simon Huttegger, Brian Skyrms, Kyle Stanford, and Charlotte Werndl for helpful comments and suggestions. This paper is an extension of my poster contribution to the 7th Annual Southern California Systems Biology Conference, January 28, 2017 at the University of California, Irvine and a talk at the SOPhiA 2017 Salzburg Conference for Young Analytic Philosophy, September 14, 2017.

References

Alon, U. (2007). *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Chapman & Hall/CRC.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., and Sherlock, G. (2000). Gene ontology: tool for the unification of biology. *Nature Genetics*, 25:25–29. doi:10.1038/75556.

Boogerd, F. C., Bruggeman, F. J., Hofmeyr, J.-H. S., and Westerhoff, H. V. (2007). Towards philosophical foundations of systems biology: introduction. In Boogerd, F. C., Bruggeman, F. J., Hofmeyr, J.-H. S., and Westerhoff, H. V., editors, *Systems Biology: Philosophical Foundations*, pages 3–19. Elsevier.

Braillard, P.-A. (2015). Prospect and limits of explaining biological systems in engineering terms. In Braillard, P.-A. and Malaterre, C., editors, *Explanation in Biology: An Enquiry into the Diversity of Explanatory Patterns in the Life Sciences*, pages 319–344. Springer.

Hucka, M., Finney, A., Sauro, H. M., Bolouri, H., Doyle, J. C., Kitano, H., Arkin, A. P., Bornstein, B. J., Bray, D., Cornish-Bowden, A., Cuellar, A. A., Dronov, S., Gilles, E. D., Ginkel, M., Gor, V., Goryanin, I. I., Hedley, W. J., Hodgman, T. C., Hofmeyr, J.-H., Hunter, P. J., Juty, N. S., Kasberger, J. L., Kremling, A., Kummer, U., Le Novère, N., Loew, L. M., Lucio, D., Mendes, P., Minch, E., Mjolsness, E. D., Nakayama, Y., Nelson, M. R., Nielsen, P. F., Sakurada, T., Schaff, J. C., Shapiro, B. E., Shimizu, T. S., Spence, H. D., Stelling, J., Takahashi, K., Tomita, M., Wagner, J., and Wang, J. (2003). The systems biology markup language (sbml): a medium for representation and exchange of biochemical network models. *Bioinformatics*, 19(4):524–531. doi:10.1093/bioinformatics/btg015.

Lesk, A. M. (2000). The unreasonable effectiveness of mathematics in molecular biology. *The Mathematical Intelligencer*, 22(2):28–37. doi:10.1007/BF03025372.

May, R. M. (2004). Uses and abuses of mathematics in biology. *Science*, 303(5659):790–793. doi:10.1126/science.1094442.

Schuster, P. (2011). Is there a Newton of the blade of grass? The complex relation between mathematics, physics, and biology. *Complexity*, 16(6):5–9. doi:10.1002/cplx.20381.

Schwartz, J. T. (1992). The pernicious influence of mathematics on science. In Kac, M., Rota, G.-C., and Schwartz, J. T., editors, *Discrete Thoughts: Essays on Mathematics, Science, and Philosophy*, pages 19–25. Springer.

Westerhoff, H. V. and Kell, D. B. (2007). The methodologies of systems biology. In Boogerd, F. C., Bruggeman, F. J., Hofmeyr, J.-H. S., and Westerhoff, H. V., editors, *Systems Biology: Philosophical Foundations*, pages 23–70. Elsevier.

Wigner, E. P. (1960). The unreasonable effectiveness of mathematics in the natural sciences. *Communications on Pure and Applied Mathematics*, 13(1):1–14. doi:10.1002/cpa.3160130102.

Wolkenhauer, O. and Ullah, M. (2007). All models are wrong ... some more than others. In Boogerd, F. C., Bruggeman, F. J., Hofmeyr, J.-H. S., and Westerhoff, H. V., editors, *Systems Biology: Philosophical Foundations*, pages 163–179. Elsevier.